TOTAL DC INTEGRATED DATA CENTERS

Peter Gross, CEO & CTO, EYP MCF Inc., USA &
Kfir L. Godrich, Sr.EE, EYP MCF Inc., USA

Abstract

Recent technological advances have made the concept of *TOTALLY DC INTEGRATED DATA CENTERS* a viable solution for the high-availability community. And the solution is here just in time – Primen estimated in a research for EPRI that by the year 2010, more than 25% of all mission-critical facilities (MCFs) will require operational availability (A_o) in the range of nine '9's (99.9999999%) [1,3]. Implementation of the *TOTALLY DC INTEGRATED DATA CENTERS* concept will make this achievable.

The unique approach we will present in this paper involves multiple data center spaces with backup from an array of three-dimensional (3D) sources or storage devices. This approach features a data center architectural design that significantly improves A_o; not just for one or two subsystems, but for the entire data center. We would like to emphasize the fact that different dimensions are based on different technologies, adding another best practice beyond the more traditional redundancy level approaches such as dual path or fault tolerance. The concept presented herein is known as $(DC)^{2\,TM}$ (pronounced 'DC Square').

1 Introduction

1.1 History

In 1997, the Intelec Advisory Committee at the 1997 Intelec conference in Melbourne, Australia, called for studies regarding the option of powering computers using DC Power Plants. In response, a white paper was developed and presented [2]. The white paper presented a very clear vision, and concluded: *"We urge all end user and specifiers of Internet and data products to purchase and install DC fed equipment as a first choice. Such an approach will significantly increase the overall reliability."*

Well, during the "Telecom Boom" it didn't happened; in fact, it took years to bring the manufacturers to grasp facts the Telecom industry figured out long ago. It also required the right business drives:

    a. The efficiency drive
    b. The High Density (HD) drive

1.2 The Efficiency Drive

In most of the cases, erecting 7X24 facilities means creating an environment for critical systems with a high Availability infrastructure, so that the risk related to mission critical supporting equipment is minimized. Unfortunately, when the relatively high cost of operating mission critical buildings is analysed, the surprising conclusion is that as much as half of the overall expenses are for the energy bill. This finding

launched additional studies on how to cut the energy costs of operating high availability MCFs. One solution favoured by the Leadership in Energy and Environmental Design (LEED) was the implementation of "Green Building" techniques in elements of the design. This solution could lead to some savings, but achieving major results is very difficult without, for example, having office buildings associated with the critical facility.

At data centres, one cost-saving measure being explored is data server efficiency. Servers are high on the list of data centre energy consumers, both directly (electrical power) and indirectly (HVAC, CRAC, CRAH units to cool the servers). See figure 1: 1U Server Power Consumption. To lower server costs, there must be improvement in the efficiency of Switched Mode Power Supplies (SMPS) at the input of the servers. Typical SMPS efficiency in active mode is between 60% and 75%; a bit lower in other modes.
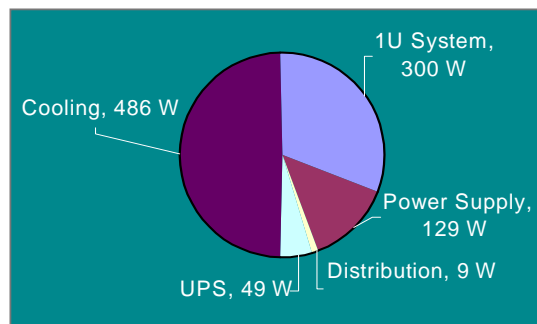

FIGURE 1: 1U SERVER POWER CONSUMPTION
(SOURCE – INTEL)

Having determined that the ultimate server board uses DC power through a DC/DC multi-output switcher and/or DC/DC converters, the most efficient solution might be supplying DC power directly to the server. This direct application of DC power will definitely improve existing energy consumption - in order to supply the power required by a 300W server, 973W of power is required. The most natural DC option must be –48VDC.

1.3 The High Density (HD) Drive

According to a recent survey [2] by Gartner Inc., blades are the fastest-growing segment of the server market. In 2004 alone, this was a US $40 billion market. The move in this direction is clear, and although the blades are currently only 3% of the market, during each and every IT upgrade today, the "blading" option is seriously considered, if not implemented.

One thing is certain – while taking out the rectifying stage in the SMPS from the servers (xU or blade's enclosure), there is suddenly additional volume leading to the same outcome – higher power density in the cabinets. If a typical AC-fed cabinet/rack on the Data Centre floor is in the range of 4-6KW, the new HD cabinets/racks are in the range of 12-15KW. Industry predictions are to increase up to 40KW per cabinet/racks in the next few years. See figure 2: Cabinets with 1U Servers (A) vs. HD Blade Servers.

The implications of this are quite profound because the requirements for the Data Centre spaces suddenly escalate to power densities up to 400W/sf. The HD solution creates new challenges, not only for powering those environments, but for cooling them as well. The weight of this discovery was fully appreciated by organizations such as ASHRAE, which lately has published updates to the previous industry assumptions. See Chart 1: Chart 1 - New Datacom Equipment Power Trend Chart - Compute And Storage Servers Split (ASHRAE). [3].

2 Today's Data Centres

2.1 Typical Configuration

Let us consider a typical, practical Tier IV Data Centre configuration. See Figure 3: Typical AC Tier IV Data Centre Module (1,000KW). Assume the critical load in figure 3 to be about 1,000KW. This could be a Data Centre module having several MW of critical load. Assume also the following configurations:

    a.   UPS systems - 2(N+1) configuration.
    b.   Engine/Generators - N+1configuration.
    c.   Utility – 2N configuration.
    d.   Mechanical loads – 2N distribution.

    e.   Dual cord equipment.Note: Such a design might also include Automatic Static Transfer Switches (ASTS) in over-crossed configuration, downstream the Power Distribution Units (PDU's). The critical load feeding chain is taking in Normal Mode the AC power of the Utility CR/A/B through an array of static double conversion UPS system. The autonomy of each module might be ten minutes or more. In the event the utility power is unavailable or out of the specifications, the engine/generators will start and supply the facility's loads. During the short period the engine/generators are starting, UPS systems are supplying the critical loads, in most cases using chemical storage.

Hybrid configuration might include High Speed Flywheels (HSF) as a single source of storage, or as a means toward improving battery systems performance at the DC-link level in the UPS. Bridging time of the HSF is inversely proportional to the power operating point. For more autonomy, HSF might be configured in parallel.

2.2 Intermediary Situations

2.2.1 AC Facility, DC Servers

So the servers are finally at –48VDC, but the infrastructure is AC. Some solutions used by several manufacturers at the cabinet level to deal with the –48VDC direct fed in HD cabinets, is to include a cluster of rectifiers at the cabinet level (bottom, top or blades chassis level).

Figure 4 is an illustration of that example - two AC inputs supply an A/B array of rectifiers in a 2(N+1) configuration.

The upper line of rectifiers is shown feeding the lower half of the cabinet, while the lower line of rectifiers is feeding the upper half of the cabinet. Some
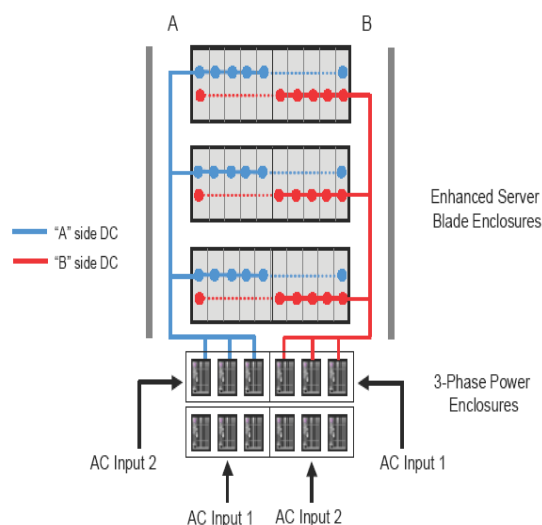

FIGURE 4: EXAMPLE OF BLADE SERVERS INTERNAL DISTRIBUTION

might express concern regarding the optimality of this solution. Basically, the same equipment remains in the cabinet, only re-organized. Another disadvantage is using relatively low power equipment. Power Electronics reliability is higher for equipment over 100KVA, also the paralleling of multiple elements yields lower availability.

Other solutions come from the Telecom world, the centralized and the distributed DC Power Plants (DCPP).

The distributed DCPP will include Rectifier/Charger (RC) cabinets at the ends of the equipment rows. This solution isn't suitable for the HD spaces because the typical maximum cabinet are in the range of 60KW. Adding multiple cabinets per side of the equipment row will be unacceptable by the IT people.

The centralized DCPP might be very reliable, although isn't practical over 500KW due to the limitation of various equipments and distribution complexity.

2.2.2   AC Facility, DC Servers

Another situation is worthy of examination - in central offices and other Telecom applications, the lack of DC servers over the years required an additional intermediary solution.

With existing –48VDC infrastructures the introduction of inverters became more and more popular.

The evolution of these inverters from 5KVA cabinets through 10KVA SMPS modules, and three phase output.

The most recent is Digital Signal Processor inverters configured in very dense configurations.

See Figure 5: Up To 15KVA Inverter packed in 7U.

These units might include 15KVA in 7U without bypass, and they might also include a static bypass unit then the output power is 12.5KVA.

FIG. 5 – UP TO 15KVA INVERTER PACKED IN 7U

3   Tomorrow's Data Centres

3.1   General

The $(DC)^{2\ TM}$ (pronounced "DC square") is a new concept in the design of Data Centres. This unique approach suits multiple Data Centres spaces with backup from an array of tri-dimensional (3D) optional sources/storage devices. The approach includes references to the architectural design of the Data Centre such that the Operational Availability is improved at the level of the whole application and not only for several sub-systems. It should be emphasized that different dimensions are based on different technologies adding another best practice besides the more traditional redundancy levels implemented, as dual path or fault tolerance.

Based on the fact that today more and more servers and peripheral equipment have direct DC power supplies, the concept of DC Data Centres or $(DC)^{2\ TM}$ is inevitable. Nonetheless, high densities of Data Centres caused by the new equipment technologies are emerging taking the conservative densities to up to 10 times higher power densities.

3.1.1   The $(DC)^{2\ TM}$ Concept

What follows is a description of the various functions within the $(DC)^2$ ™ concept – see Figure 6.

There are two main groups of relevant AC loads in the Data Centre:

> CR – critical loads are the main loads of the facility, and include all equipment installed in cabinets (servers or equivalent), on racks (telecom equipment such as routers and switches) or directly on the floor (heavy storage devices).

> MECH – mechanical loads include HVAC and all related systems, such as plumbing other essential loads supporting the critical loads.

The Utility lines feeding the Data Centre are usually Medium Voltage (MV) lines in 2N redundant configuration, Such that, a dual paths is created to the module (Data-hole) level that includes step-down transformers to bring the voltage to required level in the range of 400-660VAC, at either 50 or 60Hz.

In the event the utility lines fail, the distributed energy resources (DER) start and supply the facility loads. The DER might be an engine/generator (EG) combination, which will start in about 8 seconds (per unit) upon command, or may be integrated with main storage (MS) devices to create a continuous power supply (CPS). It also might be a source of alternative energy. MS devices include a variation of solutions based on mass or velocity (low or high speed). The output voltage of this stage is 400-660VAC, 50Hz or 60Hz. The

autonomy time may be about 15 seconds per side at full load (with a requirement to ensure the rotary machine is "walking-in" properly) – the EG or CPS will continue to supply the loads based on the amount of fuel provided to the site.

The next level downstream is the Rectifiers/Chargers (RC) level. The design includes an array of RC units in a 2(N+1) configuration. For a typical Data-hole of 1MW CR load, 3 RC units of 500KW are used. The RC units are taking the AC power and rectify it to a level of 500-550VDC. If in the level downstream of the RC units the Intermediary Storage (INTS) is using chemical battery systems (BAT/A/B), then charging functions are required when specifying the controls of the RC units. If other storage devices are in use at the INTS level, such as High Speed flywheels (HSF/A/B) then no charging functions are required. Autonomy time is about 15 seconds per side at full load, but if chemical storage is used then 10 minutes per side is typical. A hybrid approach is highly recommended, such that at one bus the storage source will be chemical while at the other HSF, therefore common technology failures are avoided.

The INTS devices are connected to the Low Voltage DC link of the system, at the Distribution Busbar System (DBBS/A/B). The DBBS voltage is in the range of 500-550VDC, distributing along the cabinets/racks pods areas.

Downstream from the DBBS/A/B for each cabinet row is a tap-off feeding a DC/DC converter on each side. The DC/DC converters are in the range of 100-150KW dropping the voltage from Low Voltage to Very Low Voltage Level of –48VDC. *Those units are the principal breakthrough making this concept viable*. The –48VDC is distributed to the cabinets row using a busbar (VLVBB/A/B) or a Remote Power Panel (RPP), which can be integrated in the DC/DC converter so that cabling systems are distributed to each cabinet in the row.

At the cabinet level, a dual path Power Distribution Frame (AB/PDF) is used to internally distribute the power. The AB/PDF's are fed from the RPP's on the end of each row, or by using tap-offs from the –48VDC busbar system.

The AB/PDF's might have a Local Storage (LS) device for several seconds only at full load to achieve dynamic stability of supercritical loads during upstream switching processes. Such kinds of LS are ultra/super-capacitor based and might be apart or integrated in the AB/PDF's.

Note: As of late, fuel cell "generators" are emerging as a recommendation of the industry, which could become a solution for lower "green" densities. If selected, this could mandate the direct supply of hydrogen lines from central sources (gas and reformers) or local bottle arrays.

4 Reliability Aspects

We analyzed the configurations presented herein supplying the same critical load of 1MW. We considered the same redundancy rules for both AC and DC configurations in our models. The results of the simulations are presented in Table 1 hereafter. As reflected from the simulations, it can be concluded that the $(DC)^{2\,TM}$ concept performance is better.

5 Energy Efficiency

As presented in the Introduction, the AC servers have relatively low efficiency. Let us look now to the AC Data Centre Distribution versus the DC Data Centre Distribution efficiencies.

Our analysis yielded the followings efficiency ranges for the configurations discussed in this paper, along with their variations (computations in the AC case included the chain between the UPS input down to the servers inputs, while in the DC case from the RC inputs down to the servers inputs):

    a. $82.0\% < \eta_{AC} < 88.0\%$
    b. $84.9\% < \eta_{DC} < 91.2\%$

Note that the use of diodes will negatively affect DC system efficiency, while the same occurs using ASTS units in the AC application.

Cooling is a major factor in the overall efficiency numbers. *In most cases the power required to <u>cool</u> the server is almost as great as the power <u>consumed</u> by the server.* Major reductions in the cooling power required can be achieved, although in the high-density environment, using water-cooled cabinets to reduce the reheating of the Data Centre spaces requires lower power for dehumidification. *Even when using relatively modest assumptions, eliminating the rectifying stage in the server SMPS and implementing the $(DC)^{2\,TM}$ design with the related cooling solution the end to end efficiency results are:*
    a.   AC system - 59.5%
    b. $(DC)^{2\,TM}$ concept – 75.2%
There is a 21% improvement in the total energy consumption.

6 Conclusions

In the future, higher availability requirements will combine with the application of newer high-density components and efficiency improvements imposed by the need to reduce Data Centre operational costs, all of which will force new design approaches.

We presented herein just such an approach –
the $(DC)^{2\,TM}$ concept, which will improve Operational Availability at the Data Centre – taking it to unprecedented levels., and in doing so provide an ingen-

ious solution for the HD spaces to yield major energy savings at the facility level.

7 References

[1] Spectrum, IEEE, 04/2005
[2] Akerlund, J. and all, -48VDC Computer Equipment Topology – An Emerging Technology, Intelec, 1998
[3] EPRI, PRIMEN - various publications.
[3] Gross P.: Needed: New Metrics, Energy User News, 08/2002
[4] Schmidt R., Beaty D.: ASHRAE Committee Formed to Establish Thermal Guidelines for Datacom Facilities
[5] Godrich, K.: Parameterization of Powering Solutions for Telecom/Datacom Clients, Intelec, 2002
[6] *(DC)* $^{2 TM}$ - US Patent pending 2005

Option	MTBF	Inherent Availability	5 Years of Operation			
			Probability of Failure	Reliability	Expected No. Failures	Total Downtime
AC Configuration	148,627.5	0.99999263	17.36%	0.8264471	0.19	0.34
DC Configuration	207,404.8	0.99999267	14.52%	0.8547904	0.16	0.31

Table 1: Reliability Analysis results

(A)   (B)

FIGURE 2: CABINETS WITH 1U SERVERS (A)
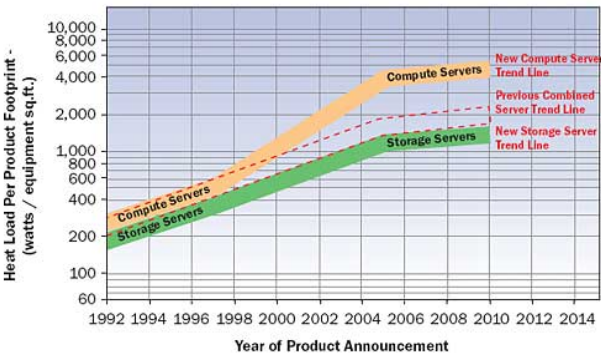VERSUS HD BLADE CABINETS (B)

CHART 1 - NEW DATACOM EQUIPMENT POWER TREND CHART - COMPUTE AND STORAGE SERVERS SPLIT (ASHRAE).

FIGURE 3: TYPICAL AC TIER IV DATA CENTER MODULE (1,000KW)

FIGURE 6: THE (DC)^2 ™ CONCEPT